



TCGTCAGGATATGCTTCCAGCCTACATGATCGATCCCGGATGATTATATAGCTGATGATG
 TTCTCTATATGTCGCAATATGCTATATACACACAGATGTCGCGATAGCATGATGATG
 TCGCATGATGCTTCT
 CACAGACTGTCGCT

Nucleotide

PubMed	Nucleotide	Protein	Genome	Structure	PopSet	Taxonomy	OMIM	Books
Search Nucleotide <input type="text"/> for <input type="text"/>							Go	Clear
Limits		Preview/Index		History		Clipboard		Details
Display	default	<input type="text"/>	Save	Text	Add to Clipboard	Get Subsequence		

☐ 1: AE000261. Escherichia coli ...[gi:1787945]

Links

LOCUS AE000261 11033 bp DNA linear BCT 01-DEC-2000
 DEFINITION Escherichia coli K12 MG1655 section 151 of 400 of the complete genome.

ACCESSION AE000261 U00096
 VERSION AE000261.1 GI:1787945

KEYWORDS .

SOURCE Escherichia coli K12.

ORGANISM Escherichia coli K12

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE 1 (bases 1 to 11033)

AUTHORS Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y.

TITLE The complete genome sequence of Escherichia coli K-12

JOURNAL Science 277 (5331), 1453-1474 (1997)

MEDLINE 97426617

PUBMED 9278503

REFERENCE 2 (bases 1 to 11033)

AUTHORS Blattner, F.R.

TITLE Direct Submission

JOURNAL Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA. Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459

REFERENCE 3 (bases 1 to 11033)

AUTHORS Blattner, F.R.

TITLE Direct Submission

JOURNAL Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA. Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459

REFERENCE 4 (bases 1 to 11033)

AUTHORS Plunkett, G. III.

TITLE Direct Submission

JOURNAL Submitted (13-OCT-1998) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA

COMMENT This sequence was determined by the E. coli Genome Project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by NIH grants HG00301 and HG01428 (from the Human Genome Project and NCHGR). The entire sequence was independently determined from E. coli K12 strain MG1655. Predicted open reading frames were determined using GeneMark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 [e-mail: mark@amher.gatech.edu]. Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible (<http://cgsc.biology.yale.edu>). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful

*admitted prior art
 page 10 concerning
 Fig. 6*

by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (<http://www.genetics.wisc.edu>). *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.

FEATURES	Location/Qualifiers
<u>source</u>	1..11033 /organism="Escherichia coli K12" /strain="K12" /sub_strain="MG1655" /db_xref="taxon:83333"
<u>protein bind</u>	<1..2 /note="central position to predicted promoter: -44.5" /bound_moiety="RhaS predicted site"
<u>promoter</u>	7..33 /note="factor Sigma70; promoter sodB; documented +1 at1733347"
<u>gene</u>	94..675 /gene="sodB" /note="b1656"
<u>CDS</u>	94..675 /gene="sodB" /EC_number="1.15.1.1" /function="enzyme; Detoxification" /note="o193; 100 pct identical to SODF_ECOLI SW: P09157 but includes initiator met; CG Site No. 15256; alternate name sodF" /codon_start=1 /transl_table=11 /product="superoxide dismutase, iron" /protein_id="AAC74728.1" /db_xref="GI:1787946" /translation="MSFELPALPYAKDALAPHISAETIEYHYGKHHQTYVTNLNNLIK GTAFEGKSLEEII RSSEGGVFNNAAQVWNHTFYWNCLAPNAGGEPTGKVAEAI AASFG SFADFKAQFTDAAIKNFGSGWTWLVKNSDGKLAIVSTS NAGTPLTTDATPLLTVDVWE HAYYIDYRNARPGYLEHFWALVNWFEVAKNLAA"
<u>repeat region</u>	741..772 /note="REP (repetitive extragenic palindromic) element; contains 1 REP sequence"
<u>gene</u>	complement(837..2006) /gene="b1657"
<u>CDS</u>	complement(837..2006) /gene="b1657" /function="putative transport; Not classified" /note="f389; residues 10-254 are 34 pct identical to aa8-252 from ARAJ_ECOLI SW: P23910 (394 aa)" /codon_start=1 /transl_table=11 /product="putative transport protein" /protein_id="AAC74729.1" /db_xref="GI:1787947" /translation="MKINYPLLALAI GAFGIGTTEFSPMGLLPVIARGVDVSI PAAGM LISAYAVGVMVGAPLMTLLLSHRARRSALIFLMAIFTLGNVLSAIAPDYM TMLLSRIL TSLNHGAFFGLGSVVAASVVPKHKQASAVATMFMGLTLANIGGVPAATWLG ETIGWRM SFLATAGLGVISMVSLFFSLPKGGAGARPEVKKELAVLMRPQVLSALLTTVLGAGAMF TLYTYISPVLQSITHATPVFVTAMLVLIGVGFSIGNYLGKGLADRSVNGTLKGFLLLL MVIMLAIPFLARNEFGAAISMVVGAA TFAVVPPLQMRVMRVASEAPGLSSSVNIGAF

promoter NLGNALGAAAGGAVISAGLGYSFVPMGAIVAGLALLLVFMSARKQPETVCVANS"
complement(2099..2127)
/note="factor Sigma70; predicted +1 start at 1735400"
promoter 2367..2395
/note="factor Sigma70; promoter purR; documented +1
at1735713"
protein bind 2500..2516
/note="central position to purR promoter:103.5"
/bound moiety="PurR documented site"
gene 2560..3585
/gene="purR"
/note="b1658"
CDS 2560..3585
/gene="purR"
/function="regulator; Purine ribonucleotide biosynthesis"
/note="o341; 100 pct identical to PURR_ECOLI SW: P15039;
CG Site No. 17989"
/codon_start=1
/transl_table=11
/product="transcriptional repressor for pur regulon, glyA,
glnB, prsA, speA"
/protein_id="AAC74730.1"
/db_xref="GI:1787948"
/translation="MATIKDVAKRANVSTTTVSHVINKTRFVAEETRNAVWAAIKELH
YSPSAVARSLKVNHTKSIGLLATSSEAAYFAEII EAVEKNC FQKG YTLILGNAWN NLE
KQRAYLSMMAQKRVDGLLVMCSEYPEPLLAMLEEYRHI PMVVM DWGEAKADFTDAVID
NAFEGGYMAGRYLIERGHREIGVIPGLERNTGAGRLAGFMKAMEEAMIKVPESWIVQ
GDFEPESGYRAMQQILSQPHRPTAVFCGGDIMAMGALCAADEMGLRVPQDVSLIGYDN
VRNARYFTPALTTIHQPKDSLGETAFNMLLDRI VNKREEPQSIEVHPRLIERRSVADG
PFRDYRR"
gene complement(3582..4514)
/gene="ydhB"
/note="b1659"
CDS complement(3582..4514)
/gene="ydhB"
/function="putative regulator; Not classified"
/note="f310; 100 pct identical to YDHB_ECOLI SW: P37598"
/codon_start=1
/transl_table=11
/product="putative transcriptional regulator LYSR-type"
/protein_id="AAC74731.1"
/db_xref="GI:1787949"
/translation="MWSEYSLEVVD A VARNGSF SAAAQELHRVPSAVSYTVRQLEEWL
AVPLFERRHRDVELTAAGAWFLKEGRSVVKMQITRQQCQ QIANGWRGQLAIAVDNIV
RPERTRQMI VDFYRHFD DVELLVFQEVFNGVWDALSDGRVELAIGATRAIPVGGRYAF
RDMGMLSWSCV VASHHPLALMDGPFSDDTLRNWPSLVREDTSRTL PKRITWLLDNQKR
VVVPDWESSATCISAGLCIGMVPTHFAKPWLN EGKWVALELENPF PDSACCLTWQQND
MSPALTWLL EYLG DSETLNKEWLREPEETPATGD"
promoter complement(4550..4578)
/note="factor Sigma70; predicted +1 start at 1737851"
promoter 4558..4586
/note="factor Sigma70; predicted +1 start at 1737901"
gene 4627..5838
/gene="ydhC"
/note="b1660"
CDS 4627..5838
/gene="ydhC"
/function="putative transport; Not classified"
/note="o403; This 403 aa ORF is 90 pct identical (7 gaps)
to 383 residues of a 400 aa protein YDHC_ECOLI SW: P37597"
/codon_start=1
/transl_table=11
/product="putative transport protein"
/protein_id="AAC74732.1"
/db_xref="GI:1787950"

/translation="MQPGKRFLVWLAGLSVLGFLATDMYLPAAAIQADLQTPASAVS
ASLSLFLAGFAAAQLLWGPLSDRYGRKPVLIGLTI FALGSLGMLWVENAATLLVLR
VQAVGVCAAAVIWQALVTDYYPQKVNRIFAAIMPLVGLSPALAPLLGSWLLVHFSWQ
AIFATLFAITVVLILPIFWLKPTTKARNNSQDGLTFTDLLRSKTYRGNVLIYAACSAS
FFAWLTGSPFILSEMGYSPAVIGLSYVPQTIAFLIGGYGCRAALQKQWQKQLLPWLLV
LFAVSVIATWAAGFISHVSLVEILIPFCVMAIANGAIYPIVVAQALRPFPATGRAAA
LQNTLQLGLCFLASLVSWLISISTPLTTTTSVMLSTVMLVALGYMMQRCEEVGCQNH
GNAEVAHSESH"
protein bind 4629..4645
/gene="ydhC"
/note="central position to predicted promoter:45.5"
/bound moiety="RhaS predicted site"
promoter 6008..6037
/note="factor Sigma70; predicted +1 start at 1739352"
gene 6129..7277
/gene="cfa"
/note="b1661"
CDS 6129..7277
/gene="cfa"
/EC_number="2.1.1.79"
/function="enzyme; Fatty acid and phosphatidic acid
biosynthesis"
/note="o382; 100 pct identical to CFA_ECOLI SW: P30010; CG
Site No. 10810; alternate name cdfA"
/codon_start=1
/transl_table=11
/product="cyclopropane fatty acyl phospholipid synthase"
/protein_id="AAC74733.1"
/db_xref="GI:1787951"
/translation="MSSSCIEEVSVPPDDNWYRIANELLSRAGIAINGSAPADIRVKNP
DFFKRVLQEGSLGLGESYMDGWECDRDLMFFSKVL RAGLENQLPHHFKDTRLIAGAR
LFNLQSKKRAWIVGKEHYDLGNLDFSRMLDPFMQYSCAYWKDADNLESAQQA KLKMIC
EKLQ LKPGMRVLDIGCGWGGLAHYMASNYDVS VVGVTISAEQQKMAQERCEGLDVTIL
LQDYRDLNDQFDRIVSVGMFEHVGPKNYDTYFAVVDRLNKPEGIFLLHTIGSKKTDLN
VDPWINKYIFPNGCLPSVRQIAQSSEPHFVMEDWHNFGADYDTTLMAWYERFLAAWPE
IADNYSERFKRMFTYYLNACAGAFRARDIQLWQVVF SRGVENGLRVAR"
gene complement(7317..7958)
/gene="ribE"
/note="b1662"
CDS complement(7317..7958)
/gene="ribE"
/EC_number="2.5.1.9"
/function="enzyme; Biosynthesis of cofactors, carriers:
Riboflavin"
/note="f213; 100 pct identical to RISA_ECOLI SW: P29015;
CG Site No. 11923"
/codon_start=1
/transl_table=11
/product="riboflavin synthase, alpha chain"
/protein_id="AAC74734.1"
/db_xref="GI:1787952"
/translation="MFTGIVQGTAKLVSIDEKPNFRTHVVELPDHMLDGLGTGASVAH
NGCCLTVTEINGNHVSFDLMKETLRITNLGDLKVGDWVNVERAAKFSDEIGGHLMSGH
IMTTAEVAKILTSENNRQIWFKVQDSQLMKYILYKGFIDGISLTVGEVTPTRFCVH
LIPETLERTTLGKKKLGARVNIEIDPQTQAVVDTV ERVLAARENAMNQP GTEA"
promoter complement(7998..8026)
/note="factor Sigma70; predicted +1 start at 1741299"
promoter 8111..8140
/note="factor Sigma70; predicted +1 start at 1741455"
gene 8173..9546
/gene="ydhE"
/note="b1663"
CDS 8173..9546
/gene="ydhE"
/function="putative transport; Not classified"

/note="o457; 99 pct identical to fragment YDHE_ECOLI
SW:P37340 (230 aa) but contains 227 additional C-ter
residues"
/codon_start=1
/transl_table=11
/product="putative transport protein"
/protein_id="AAC74735.1"
/db_xref="GI:1787953"
/translation="MQKYISEARLLLALAI PVILAQIAQTAMGFVSTVMAGGYSATDM
AAVAIGTSIWLPAILFGHGLLLALTPVIAQLNGSGRRERIAHQVRQGFVLAGFVSVLI
MLVLWNAGYIIRSMENIDPALADKAVGYLRALLWGAPGYLFFQVARNQCEGLAKTKPG
MVMGFIGLLVNI PVNYIFIYGHFGMPGLGGVGCVATAAVYWVFLAMVSYIKRASM
RDIRNEKGTAKPDPAVMKRLIQLGLPIALALFFEVTLFAVVALLVSPLGIVDVAGHQI
ALNFSSLMFVLPMSLAAAVTIRVGYRLGQGSLDAQTAARTGLMVGVCMTLTAIFTV
SLREQIALLYNDNPEVVTLAHLMLLAAVYQISDSIQVIGSGILRGYKDRSIFYITF
TAYWVLGLPSGYILALTDLVPEPMGPAGFWIGFII GLTSAAIMMMLMRFLQRLPSAI
ILQRASR"

gene complement(9587..10843)

/gene="b1664"

CDS complement(9587..10843)

/gene="b1664"

/function="putative enzyme; Not classified"

/note="f418; This 418 aa ORF is 31 pct identical (50 gaps)
to 391 residues of an approx. 864 aa protein YEJO_ECOLI
SW: P33924"

/codon_start=1

/transl_table=11

/product="possible enzyme"

/protein_id="AAC74736.1"

/db_xref="GI:1787954"

/translation="MGSDAKNLMSDGNVQIVKTGEVIGATQLTEGELIVEAGGRAENT
VVTGAGWLKVATGGIAKCTQYGNNGTLSVSDGAIATDIVQSEGGAI SLSTLATVNGRH
PEGEFSVDKGYACGLLLENGGNLRVLEGHRAEKIILDQEGGLLVNGTSAVVVDEGGE
LLVYPGGEASNCEINQGGVFMLAGKASDTLLAGGTMNNLGGEDSDTIVENGSIYRLGT
DGLQLYSSGKTQNL SVNVGGRAEVHAGTLENAVIQGGTVILLSPTSADENFVVEEDRA
PVELTGSVALLDGASMIIGYGAE LQQSTITVQQGGVLILDGSTVKGDSVTFIVGNINL
NGGKLWLITDAATHVQLKVKRLRGEGAICLQTSAKEISPDFINVKGEVTGDIHVEITD
ASRQTL CNALKLQPD EDGIGATLQPA"

promoter complement(10887..10914)

/note="factor Sigma70; predicted +1 start at 1744188"

promoter complement(10926..10952)

/note="factor Sigma32; predicted +1 start at 1744227"

BASE COUNT 2627 a 2894 c 2660 g 2852 t
ORIGIN

```

1  tcatcttttg  tctcaccttt  taatttgcta  ccctatccat  acgcacaata  aggctattgt
61  acgtatgcaa  attaataata  aaggagagta  gcaatgtcat  tcgaattacc  tgcactacca
121  tatgctaaag  atgctctggc  accgcacatt  tctgcggaaa  ccatcgagta  tcactacggc
181  aagcaccatc  agacttatgt  cactaacctg  aacaacctga  ttaaaggtag  cgcgtttgaa
241  ggtaaatcac  tggaagagat  tattcgcagc  tctgaagggtg  gcgtattcaa  caacgcagct
301  caggtctgga  accatacttt  ctactggaac  tgcctggcac  cgaacgccg  tggcgaaccg
361  actggaaaag  tcgctgaagc  tatcgccgca  tcttttgga  gctttgccga  tttcaaagcg
421  cagtttactg  atgcagcgat  caaaaacttt  ggttctggct  ggacctggct  ggtgaaaaac
481  agcgatggca  aactggctat  cgtttcaacc  tctaacgcgg  gtactccgct  gaccaccgat
541  gcgactccgc  tgctgaccgt  tgatgtctgg  gaacacgctt  attacatcga  ctatcgcaat
601  gcacgtcctg  gctatctgga  gcacttctgg  gcgctgggtg  actgggaatt  cgtagcgaaa
661  aatctcgctg  cataataact  gatggcaaat  gcagcattgc  ctgaagcgct  acgcttatca
721  ggctacgcg  gatcatcgat  gtaggtcgga  taaggcactc  gccgatccg  gcaagataaa
781  tcgcacgttg  tcagcaactg  taacgcagaa  ggttatcctt  ctgcgttttt  gtttaattag
841  ctgttagcaa  cgcaaaactg  ttcaggttgt  tttctggctg  acataaacac  cagcaataat
901  gccagtcccg  cgacaatcgc  tcccatcacc  ggcacaaaag  tgtatcccag  cccagcgga
961  attaccgcac  caccagcagc  tgctcccagc  gcatttccaa  gattaaaggc  accaatattg
1021  actgatgaag  acagacctgg  cgcttcactg  gcgacacgca  tcacgcgcat  ctgtaacggc
1081  ggtacgaccg  caaaggttgc  tgcgccccac  accaccatgc  taatagctgc  gccgaactca
1141  ttgcgggcca  ggaacgggat  tgccagcata  atcaccatca  acaacaacaa  aaagcctttc
1201  aacgtgccgt  taactgaacg  atctgccagt  ttgccgcca  gatagttacc  gatagagaat

```

```
1261 ccgacaccaa tcagcaccag cattgccgtg acgaacaccg gtgttgcggt ggtaatactt
1321 tgcagtaccg gagagatata ggtgtagaga gtaaaccattg caccagctcc cagtaccgtc
1381 gtcagcaatg cagacagcac ctgcggacgc attaataccg ccagctcttt ttctacttca
1441 ggctcgtgcc ctgcaccacc tttaggtaat gagaagaaca gacttaccat tgaaatcact
1501 cccagccccg ccgttgccag aaatgacatc cgccagccga tggtttcacc caaccaggtc
1561 gccgccggca cgccaccgat atttgccagg gtttaaccca taaacatagt ggcaactgcg
1621 ctggcctggt tatgttttgg caccacgctt gcggccacga ctgaacccaa accaaaaaat
1681 gtcctcgtgat tcaggctggt caaatgctg gaaagcatca gggtcatata atccggcgcg
1741 atggcggaat gtacgttgcc gagcgtgaaa attgccatca ggaaaatcaa cgcaactgcg
1801 cgggcacgat gagaaagtag aagcgtcatc agcggcgcgc caaccattac gccaaactgca
1861 taggcactga ttaacattcc ggcagcggga atcgagacat ccacaccgcg cgcaatgacg
1921 ggcaacaagc ccattggcga gaactccgtt gtcccagatac caaacgcgcc aatcgccagc
1981 gccagcaacg gatagttaat tttcatgcct tatctccacc tcttcgctgc attacgcgat
2041 attcattaaa gtggcgaaaag catgacagca atcacaaaaa aatgaaaaata acaaaaagag
2101 aaaacacttt tgccattttg ctaacaaaca ggaaggagat gcgagggaga acgcgctccc
2161 tcgagaggaa atcagtgcag cgcgcgagtc aaaccacagg ctacgatcaa accgaggacg
2221 ataatcgttg ttaccagtga aaatttaagg tcggtgctca tcaagttttc tcctttttta
2281 ttaccacaca aaaagtgata ttacgcattt ttacacactg tgatgaaaaa atctcccgtc
2341 atttataatg ataagtgttt ttaccacttc cctttttcgt caagatcggc caaaattcca
2401 cgcttacact atttgcgtac tggccattga ccccttcctg acgctccgtg tcgtttttcc
2461 ggctgaccgc aacacttttg ttgtgcgtaa ggtgtgtaaa ggcaaacggt taccttgcca
2521 ttttgaggga gctgaagtta gggctcggag tgaaatggaa tggcaacaat aaaagatgta
2581 gcgaaacgag caaacgtttc cactacaact gtgtcacacg tgatcaacaa aacacgtttc
2641 gtgcgtgaag aaacgcgcaa cgcggtgtgg gcagcgatta aagaattaca ctactccctc
2701 agcgcggtgg cgcgtagcct gaaggttaac cacaccaagt ctatcggttt gtggtcgacc
2761 agcagcgaag cggcctattd tgccgagatc attgaagcag ttgaaaaaaa ttgcttccag
2821 aaaggttaca ccctgattct gggcaatgcg tggacaatc ttgagaaaca gcgggcttat
2881 ctgtcgatga tggcgcaaaa acgcgtcgat ggtctgctgg tgatgtgttc tgagtacca
2941 gagccgttgc tggcgatgct ggaagagtat cgccatatcc caatggtggt catggactgg
3001 ggtgaagcaa aagctgactt caccgatgcg gtcattgata acgcgttcga aggcggctac
3061 atggccgggc gttatctgat tgaacgcggt caccgcgaaa tcggcgctcat ccccgcccg
3121 ctggaacgta acaccggcg caggccgctt gccggtttta tgaaggcgat ggaagaagcg
3181 atgatcaagg tgcggaaaag ctggattgtg cagggtgact ttgaacctga atccggttat
3241 cgcgccatgc agcaaatcct gtgcgagcgt catcgcccta ctgccgtctt ctggttggcg
3301 gatcatcatg caatgggcgc actttgtgct gctgatgaaa tgggcctgcg cgtcccgcag
3361 gatgtttcgc tgatcggtta tgataacgtg cgcaacgcgc gctattttac gccggcgctg
3421 accacgatcc atcagccaaa agattcgctg ggtgaaacag cgttcaacat gctgttggat
3481 cgtatcgtca acaaacgtga agaaccgcag tctattgaag tgcatccgcg cttgattgaa
3541 cgccgctccg tggctgacgg cccgttccgc gactatcgtc gttaatcacc cgttgccgga
3601 gtctcttccg gctcccgcag ccactcctta ttcagcgtct cactatcgcc gagatactca
3661 agcaaccagg ttaacgcagg cgacatacca ttttgcgtgc atgtcagaca acatgccgaa
3721 tccggaagg gggtttccag ttctaattgt acccacttcc cctcattaag ccacggtttg
3781 gcgaaatgtg ttggcaccat ccctatgcat aatcctgccg agatacaggt tgcgatgat
3841 tcccagtcag gcacgacaac tcttttttgg ttatccagca accaggtaat acgtttaggt
3901 agcgttcgcg aggtgtcttc gcgcaccaac gacggccagt tgcgcaacgt atcatcgctg
3961 aacgggccat ccatcaacgc cagcgggtgg tgactggcaa caacgcaact ccagcttagc
4021 atccccatat cccggaaggc ataacgaccg cctaccggaa tcgcgcgtgt tgcgccaatc
4081 gccagttcca cgcgcccgtc ggaaagcgca tcccagacac cgttgaacac ttcctgaaag
4141 acaagaagtt cgacatcatc aaaatggcga taaaatcaa cgatcatctg ccgtgtacgt
4201 tctggcctga caatattatc cactcgata gctaactgac cgcgccagcc gttcgtatc
4261 tgctgacatt gctggcgggt gatctgcatt ttttgacaa cagagcgccc ttctttgaga
4321 aaccacgctc cagcagcggg cagctccaca tcacggtgcc gtcgttcaaa gagcggcacc
4381 gccagccact cttccagctg acgcacggta tagctgaccg cagaaggaac gcgatgcagc
4441 tcctgtgccg cagcgctaaa actaccatta cgcgctaccg catcaacaac ttcgagttaa
4501 tattctgacc acatagtctg cctgcaaaat ttttgaaacc agtcatcaaa tattaccgtt
4561 tcacaacact aatttcactc cctacacttt gcggcggtgt ttaattgaga gatttagaga
4621 atatacatgc aacctgggaa aagattttta gtctggctgg cgggtttgag cgtactcggg
4681 tttctggcaa ccgatatgta tctgctgctg ttcgcccga tacaggccga cctgcaaacg
4741 cctgcgtctg ctgtcagtgc cagccttagt ctgttccttg ccggttttgc cgcagccag
4801 cttctgtggg ggccgctctc cgaccgttat ggtcgtaaac cggattattt aatcggcctg
4861 acaatttttg cgtaggttag tctggggatg ctgtgggtag aaaacgcgcg tacgctgctg
4921 gtattgcgtt ttgtacaggc tgtgggtgtc tgcgcccgg cggttatctg gcaagcatta
4981 gtgacagatt attatccttc acagaaagtt aaccgtattt ttgcggccat catgccgctg
5041 gtgggtctat ctccggcact ggctcctctg ttaggaagct ggctgctggt ccatttttcc
5101 tggcaggcga ttttcgccac cctgtttgcg attaccgtgg tgctgattct gcctattttc
```

```
5161 tggctcaaac ccacgacgaa ggcccgtaac aatagtcagg atggtctgac ctttaccgac
5221 ctgctacgtt ctaaaacctt tcgcggaac gtgctgatat acgcagcctg ttcagccagt
5281 ttttttgcac ggctgaccgg ttcaccgttc atccttagtg aaatgggcta cagcccggca
5341 gttattgggt taagttatgt cccgcaaact atcgcgtttc tgattgggtg ttatggctgt
5401 cgcgcgcgc tgcagaaatg gcaaggcaag cagttattac cgtggttget ggtgctgttt
5461 gctgtcagcg tcattgcgac ctgggctgcg ggcttcatta gccatgtgtc gctggtcgaa
5521 atcctgatcc cattctgtgt gatggcgatt gccaatggcg cgatctaccc tattgttgtc
5581 gcccagcgcg tgcgtccctt cccacacgca actggtcgcg ccgcagcgtt gcagaacact
5641 cttcaactgg gtctgtgctt cctcgcaagt ctggtagttt cctggctgat cagtatcagc
5701 acgccattgc tcaccaccac cagcgtgatg ttatcaacag taatgctggt cgcgctgggt
5761 tacatgatgc aacgttgtga agaagttggc tgccagaatc atggcaatgc cgaagtcgct
5821 catagcgaat cacactgatc tatatcgata tacttatact taggctgcta acaaaatttt
5881 gttgtatgat tgaaattagc ggcctatact aatttcgagt tgtaaagct acgataaata
5941 ttatgttttt acggggacag gatcggtccc gactcactat ggatagtcac ttcggcaagg
6001 gttcctcctt tccctctgtt ctacgtcgga ttatagactc gcggtttttt ctgagagatt
6061 tctcacaagg cccaaaaagc gtctacgctg ttttaagggt ctgatcaccg accagtgatg
6121 gagaaactat gaggttcatc tgtatagaag aagtcagtgt accggatgac accgtgtacc
6181 gtatcgccaa cgaattactt agccgtgccc gtatagccat taacggttct gcccgcgagg
6241 atattcgtgt gaaaaacccc gattttttta aacgcgttct gcaagaaggc tctttggggg
6301 taggcgaaag ttatatggat ggctgggtgg aatgtgaccg actggatatg ttttttagca
6361 aagtcttacg cgcaggtctc gagaaccaac tcccccatca tttcaaagac acgctgcgta
6421 ttgccggcgc tcgtctcttc aatctgcaga gtaaaaaacg tgcctggata gtcggcaaa
6481 agcattacga tttgggtaat gacttgttta gccgcagtgt tgatcccttc atgcaatatt
6541 cctgcgttta ctggaagat gccgataatc tggaaatctgc ccagcaggcg aagctcaaaa
6601 tctattgtga aaaattgcag ttaaaacagc ggatgcgcgt actggatatt ggctgcgggt
6661 ggggcgcgact ggcacactac atggcatcta attatgacgt aagcgtggtg ggcgtcacca
6721 tttctgccga acagcaaaaa atggctcagg aacgctgtga aggcctggat gtcaccattt
6781 tgctgcaaga ttatcgtgac ctgaacgacc agtttgatcg tattgtttct gtgggatgt
6841 tcgagcacgt cggaccgaaa aattacgata cctattttgc ggtggtggat cgtaatttga
6901 aaccggaagg catattcctg ctccatacta tcggttcgaa aaaaaccgat ctgaatgttg
6961 atccctggat taataaatat atttttccga acggttgctt gccctctgta cgccagattg
7021 ctcagtcagg cgaaccccac tttgtgatgg aagactggca taacttcggt gctgattacg
7081 atactacgtt gatggcgtgg tatgaacgat tctcgcgcgc atggccagaa attgcccata
7141 actatagtga acgctttaaa cgaatgttta cctattatct gaatgcctgt gcaggtcctt
7201 tccgcgcccg tgatattcag ctctggcagg tcgtgttctc acgcggtgtt gaaaacggcc
7261 ttcgagtggc tcgctaaagg ctattctatc gccccctctc cgggggcatg ttcagatcag
7321 gcttctgtgc ctggttgatt catggcattt tctcgtgccg ccagcacacg tttaccgta
7381 tctaccactg cctgagtttg tggatcgatt tcaatgttga cgcgtgcgcc aagttttttc
7441 ttcccaagag tcgtgcgttc cagtgtttcc ggaattaaat ggacgcaaaa acgctgtggc
7501 gtgacttcgc cgacggtcag gctaataccg tcgatgccaa taaatccttt gtacagaata
7561 tatttcatac actgactatc ctggacttta aaccagatct ggcgattatt tctgaggtt
7621 aatattttcg acacttcagc agtggctata atatgacctg acattaaagt tccgccattt
7681 tcatcactga atttcgccgc acgctcaacg tttacccaat cccccacttt taaatcgcca
7741 agattggtaa tgcgtaacgt ttctttcatc aggtcaaaac tgacatggtt gccgttaatt
7801 tccgtcacgg tcaggcagca accgttatgc gccacggaag caccggtttc caggccgtcc
7861 agcatgtggt cgggtaactc caccacatgc gtacgaaaat ttggtttctc gtcaatcgac
7921 accagttttg cggtgccctg tacaatcccc gtaaacatac ttacaactcc tgaaatcagt
7981 taagacattc tgttcagcac aatagcaggt ggaaaacgcc cttaccagtg aaggggtaag
8041 aatggctatt ttttactgg agaattaata aatcctcgct acaatagact gaatttcccc
8101 tgccttctct ttttgcgtcc cattcaggcg gctttttagt ctctcatata actacaaata
8161 aaagggtgtt acgtgcagaa gtatatcagt gaagcgcgtc tgttattagc attagcaatc
8221 ccggtgattc tcgcgcaaat cgcccaaact gcgatgggtt ttgtcagtac cgtgatggcg
8281 ggcggtata gtgccaccga catggcggcg gtcgctatcg gtacttctat ctggcttcgg
8341 gcgatcctct ttggtcacgg actgctgctg gcattaacgc cggttatcgc gcaattaaat
8401 gggtccgggt gacgtgagcg cattgcgcat caggtgcgac aaggtttctg gctggcaggt
8461 tttgtttccg ttctcattat gctggtgctg tggaaatgag gttacattat ccgctccatg
8521 gaaaacatcg atccggctct ggcggaacaa gccgtgggtt atctgcgtgc gttgttgtgg
8581 ggcgcgccgg gatattctgt cttccaggtt gcccgtaacc agtgtgaagg tctggcaaaa
8641 accaagccgg gtatggtaat gggctttatc ggcctgctgg tgaacatccc ggtgaactat
8701 atctttatct atggtcattt cggtatgcct gagctcggtg gcgttggttg tggagctgct
8761 actgcggcgg tgtattgggt catgttcctt gccatgggtt cttacattaa acgcgccgcg
8821 tccatgcgcg atattcgtaa cgaaaaaggc accgcaaaac ccgatcctgc gggtatgaaa
8881 cgactgattc aactcggttt gccgattgct ctggcactgt tctttgaagt gacactgttt
8941 gccgtcgtgg ctctgttagt gtctccgctc ggtattgttg atgtcgcagg acaccagatt
9001 gccctgaact ttagttcact aatgttcgtg cttccaatgt cgctggcggc agcggtaact
```

```
9061 atccgcgtag gttatcgtct gggtcagggc tcaacgctgg atgcgcaaac cgctgcgcgg
9121 accgggctta tgggtgggtgt ctgtatggca accctgacgg ccattttcac ggtttactg
9181 cgggagcaaaa tcgccctgtt gtacaacgac aatcccaggg ttgtaacgct ggctgcgcat
9241 ttgatgtttgc tggcggcggg atatcagatt tctgactcaa tccaggtgat tggcagtg
9301 attttgcgtg gttataaaga tacgcgttcc attttctata ttacctttac ggcttactgg
9361 gtgctgggct tgccaagcgg ctatattctg gcactgaccg atctgggtcg tgaacctatg
9421 gggccagcag gcttctggat aggctttatt attggcctga cgtcggcagc cattatgatg
9481 atgttgcgta tgcggttcct gcaacgtctg ccgtcagcca tcattctgca acgagcatcc
9541 cgctaataaaa gacaaggcgc aaccttcacg ggttgcgctt gtatttttac gcaggctgga
9601 gcgttgcgcc aatcccgctc tcgtctggct gtaatttcag agcgttacac agagtgtgcc
9661 gactggcatc tggtatctca acgtgtatat ccccggtaac ttcccccttc acattgatga
9721 agtcaggtga aatttctttt gcactgggtt gcaggcaaat cgctccctct ccgcgcaggc
9781 gtttactttt cagttgcaca tgcgttgccg catcagtgat cagccacagt tttccaccat
9841 tcagattgat gttaccaaca atgaaagtga cactgtcacc ttttaccgta ctgccgtcta
9901 gaatcaacac accgccctgc tgtacagtaa tcgttgattg ttgcagctct gcgccataac
9961 caataatcat tgaagcgccg tccagtaatg caacactccc ggtcagttca accggtgcgc
10021 gatcttcctc tacgacaaaa ttttcgtccg cgctgggtgg tgacaacagg atcactgtcc
10081 caccttgtat taccgcgttt tccagcgtac cggcatgcac ttacgccga ccaccacat
10141 tcacggacag gttttgcgtc ttaccggaac tgtagagctg aaggccatcc gtccccagac
10201 gataaatgga tccattctca acaatagtgt cagagtcttc accaccgaga ttattcatgg
10261 tgccaccagc aagcaacgta tcaactggct tcccggccag cataaaaacg ccgccctgat
10321 taatctcaca attgctggct tccccacctg gatacaccaa caattcacca ccttcatcta
10381 ccacgaccgc tgaggttgtc ccattaacca acaggccgcc ctcttgatcg agaatgattt
10441 tttccgcgcg atgtccttcc agtacacgca ggttaccgcc attttccagc aacaaaccgc
10501 aggcataacc tttatcaacg ctgaattcac cttcgggatg gcggccatta accgtagcga
10561 gcgtagagag actaattgcg cctccctcgg actgaacaat atctgtggca atggccat
10621 cgctgaccga tagcgtgcca ttgttaccgt actgtgtgca tttggcgatc ccaccggtt
10681 ccactttcaa ccagccagcc cccgtgacca cggatatttc ggctcttccg ccagcttcaa
10741 caatcaactc gccttcagta agttgcgtcg cgccaatgac ctgcgcggtc ttaacaattt
10801 gcacgttccc gtcgctcatc aagtttttcg catcagatcc catgatttat tcctttgctg
10861 catctgtgtg cctttattgc tacctaagtg taaaggctac ggaggattta tccacgacag
10921 atttgagatg gtggcaaaaa actctgttta aactctgata cacgaattat tgggttgtat
10981 cagatgtaaa tgcgatcctg aataaaaaatc acccttgcaa atcaacaaaa tat
```

//

Revised: July 5, 2002.

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

Oct 21 2002 11:56:56